Pergamon Press plc. 1990. Printed in the U.S.A.

62

Results From the European Multicenter Study on Lead Neurotoxicity in Children: Implications for Risk Assessment[1]

GERHARD WINNEKE,*[2] ARTHUR BROCKHAUS,† ULRICH EWERS,†
URSULA KRÄMER‡ AND MANFRED NEUF*

*Department of Psychophysiology, †Department of Chemistry and ‡Department of Biostatistics
Medizinisches Institut für Umwelthygiene an der Heinrich-Heine-Universität
Düsseldorf, D-4000 Düsseldorf, W. Germany

WINNEKE, G., A. BROCKHAUS, U. EWERS, U. KRÄMER AND M. NEUF. *Results from the European multicenter study on lead neurotoxicity in children: Implications for risk assessment.* NEUROTOXICOL TERATOL 12(5) 553–559, 1990.—In order to improve dose-response information on neurobehavioral effects of environmental lead exposure in children, the World Health Organization, Regional Office for Europe (WHO/EURO), in collaboration with the Commission of the European Communities, initiated this international study which was planned, executed and evaluated between 1984 and 1989. Eight groups from eight European countries (Bulgaria, Denmark, Greece, Hungary, Italy, Roumania, W. Germany and Yugoslavia) took part. A common study protocol with inherent quality assurance elements was developed to achieve comparability. Blood-lead concentrations (PbB) were the main markers of exposure. The WISC (4 subtests) for psychometric intelligence, the Bender Gestalt test (GFT version) and the Trail-Making test for visual-motor integration, the Vienna Reaction Device and a delayed RT task for reaction performance, and the Needleman scales for behavior ratings served as behavioral endpoints. All individual studies taken together represent a sample size of 1879 school-age children and cover a PbB range from below 5 to about 60 μg/100 ml. Overall statistical evaluation of outcome was done by multiple regression analysis using a uniform confounder model. The strongest and most consistent effects occurred for the Bender Gestalt test (GFT version) and for serial choice reaction performance (Vienna Device). The degree of association with PbB was significant for these variables, although the contribution of PbB to the observed variance never exceeded 0.8%. Psychometric intelligence was also negatively affected, although the consistency of outcome between studies was poor, and the association with PbB only borderline. An effort towards risk assessment was made by calculating the proportion of children at risk, using the observed regression coefficients as well as means and standard deviations.

Environmental lead Neurobehavior Multicenter study Children Visual-motor integration
Reaction performance Intelligence Risk assessment

REALISTIC risk assessment for lead-induced neurobehavioral deficit in environmentally exposed children requires sufficiently consistent results from independent studies as well as the documentation of dose-response contingencies; despite an impressive amount of work in this field such information is still rather limited, as pointed out in recent reviews (6, 12, 14). For this reason, the World Health Organization, Regional Office for Europe (WHO/EURO), together with the Commission of the European Communities (CEC), decided to approach this problem by organizing a collaborative multicenter study in which results from several independent cross-sectional studies should be combined on the basis of a common study protocol with inherent quality assurance elements. The main purpose of combining individual studies within a single overall analysis was to increase the sample size and to broaden the range of the exposure dimension; both these aspects are important prerequisites for the establishment of dose-response contingencies.

The WHO/CEC collaborative study on the neurobehavioral toxicity of environmental lead exposure in children was planned, performed and evaluated between 1984 and 1989. Eight laborato-

[1]Based on a paper at the conference "Methods in Behavioral Toxicology and Teratology," October 4–6, 1989 in Little Rock, AR.

[2]Requests for reprints should be addressed to Gerhard Winneke, Medizinisches Institut für Umwelthygiene an der Heinrich-Heine-Universität, Düsseldorf Auf'm Hennekamp 50, D-4000 Düsseldorf, W. Germany.

ries from eight European countries participated. Participants (country, city, principal investigator = PI) were from Bulgaria (Sofia, PI = Z. Zaprianov), Denmark (Aarhus, PI = O. Nörby-Hansen), Greece (Athens, PI = A. Hatzakis), Hungary (Budapest, PI = P. Rudnai), Italy (Modena, PI = G. Vivoli), Roumania (Bucharest, PI = M. Cucu), W. Germany (Düsseldorf, PI = G. Winneke) and Yugoslavia (Zagreb, PI = O. Weber). The "Medizinisches Institut für Umwelthygiene an der Universität Düsseldorf" (Medical Institute of Environmental Hygiene, University of Düsseldorf) served as the coordinating institute (CI).

One thousand eight hundred and seventy-nine school-aged children between age 6 and 11 years were examined. Blood-lead concentrations covered the range from below 5 to about 60 μg/100 ml. Results from some individual studies have already been published (3, 9, 15, 17). The present report is restricted to the main aspects of methodology and outcome of the overall analysis of the full set of data, and its discussion in terms of risk assessment. More comprehensive reports of the full study are being prepared (13,16).

METHOD

A common study protocol was developed covering the following aspects: age range and recommended sample size, markers of lead exposure including quality control strategies, mandatory and optional outcome measures including quality assurance efforts, recommendations for extraneous and confounding variables to be considered, criteria of exclusion, as well as crucial aspects of procedure, study design and statistical analysis. The following description covers only those methodological aspects which were basic to the final overall analysis of the full data set.

Exposure Assessment and Quality Assurance

Both blood-lead and tooth-lead concentrations were initially considered as markers of environmental lead exposure. Flameless atomic absorption spectrophotometry (AAS) was the analytical method in all laboratories; anodic stripping voltammetry was additionally used in one study. External quality control for blood-lead values was performed within the framework of the UK Quality Assurance Scheme (UKQAS) using the criteria of the Wolfson Research Laboratory (Birmingham), whereas internal quality control was performed using Behring control samples of 15, 41 and 73 μg/100 ml, and two standard deviations (SD) around the mean as the criterion of acceptance. Only five groups took part in the quality control programme for tooth-lead analysis using ground tooth powder with known lead concentrations provided by the CI in Düsseldorf. Both internal and external quality controls were performed, and criteria for acceptance were 2 SD around the mean, again.

The initial plan called for using both blood-lead and tooth-lead concentrations as markers of environmental lead exposure. However, in due course of the study, it became clear that it was difficult to collect a sufficient number of teeth for analysis. In addition, the limitations of whole tooth-lead concentrations (PbT) due to age of tooth shedding and bias of tooth giving (higher IQ, better social status) were better recognized and confirmed in this study. It was, therefore, decided to use only blood-lead levels (PbB) as markers of individual environmental lead exposure for the overall data analysis. This decision also appears justified vis a vis the observation that PbBs in children tend to correlate highly ($r>.80$) over a period of at least three years (17).

Outcome Measures and Quality Assurance

The selection of outcome measures was not primarily guided by theoretical considerations concerning the effects of low-level lead exposure on neurobehavioral functions. Instead, the guiding principle was to include measures which had been used in previous studies, or which, according to clinical experience in brain-damaged children, promised to provide additional relevant information.

Functional areas covered by the agreed upon battery of tests included psychometric intelligence [Wechsler Intelligence Scale for Children (WISC)], visual-motor integration [Bender Gestalt test and the Trail-Making test (TMT)], reaction performance [Delayed Reaction Time (DRT) and Vienna Reaction Device (VRD)], as well as general behavior as rated by parents and/or teachers by means of simple rating scales (7).

Psychometric intelligence. The Wechsler Intelligence Scale for Children (WISC) consists of a verbal and a performance part represented by 5 or 6 subscales each. Although five laboratories had used the full WISC in their individual studies, the final overall data analysis was based on only those four subscales, which had been used by all the groups. These were "Vocabulary" and "Comprehension" from the verbal part, and "Picture Completion" as well as "Block Design" from the performance part. Because of the high correlations ($r>.9$) between the mean standard score for these four subscales and the full-scale WISC score, calculated from those studies which had used the full WISC, the loss in precision by using the reduced WISC version is marginal.

Visual-motor integration. The "Trail-Making Test, form A" (TMT-A) and the "Bender Gestalt test" (GFT version) were used here. Both are established clinical tools for the detection of brain damage in children.

The TMT-A is part of the Halstead-Reitan Neuropsychological Test Battery (10). The child is asked to connect a random pattern of letters in the correct sequence as fast as possible. The time needed to finish the task is taken. This test was applied in only 5 of the 8 studies.

The "Bender Gestalt test" was given in its clinical form (1). The task of the children is to reproduce simple line drawings presented on cards. The evaluation of these reproductions in terms of accuracy was done by means of the German scoring system of the "Göttinger Formreproduktionstest, GFT" (11); raw error scores were used instead of age-adjusted standard scores, because these were developed on rather small samples. In addition to the normal version (GFT-N) which requires the reproductions to be drawn on clean white paper, the background interference version (2) was used as well (GFT-I). Here, the reproductions are to be drawn on paper covered with wavy lines; the underlying idea is to interfere with perceptual figure-ground separation. Clinical experience has shown that brain-damaged children exhibit more pronounced performance deficit with the GFT-I than with the normal version (19).

Reaction performance. A delayed reaction time test (DRT) and a serial choice reaction test, namely the "Wiener Determinationsgerät" [Vienna Reaction Device (VRD)] were marked as optional in the study protocol. The DRT was used in only three and the VRD in six of the studies.

The DRT was developed at the Institute of Psychiatry (London) by Yule and co-workers (4) following the original description given by Needleman *et al.* (7). In this computerized version the child must respond to a visual signal following a preparatory auditory signal after 3 or 12 seconds delay, respectively. The VRD is a serial choice reaction apparatus (Schuhfried, Enzensdorf, Austria) in which the children have to respond to auditory (high and low pitch) and coloured light signals by pressing the appropriate response buttons (5). Both hits and errors were recorded at two conditions of signal rate, a slow and a fast one, respectively (18). From the six studies two were disregarded for methodological uncertainties, namely change of task difficulty between age levels.

General behavior ratings. Both teachers and parents were asked to rate the children's behavior in the classroom or at home using the items originally described by Needleman *et al.* (7). These are ten simple yes/no items covering observable aspects of behavior (e.g., is this child easily distracted during his/her work? is he/she easily frustrated by difficulties? can he/she follow simple directions?, etc.); these items will hereafter be referred to as the Needleman scales. A principal components analysis was done for the ten teacher scales for each study. The single items were scored such that higher values indicate poorer behavior, and the mean of 6 items constituting the main component in each study group was taken as the outcome variable. Parent ratings were treated likewise.

Quality assurance efforts for behavioral measures. The following steps were taken to improve the comparability of the psychological test data: 1) key aspects of the testing procedure were standardized in a detailed manner before starting the main study; 2) the behavioral part of each study was to be supervised by a senior psychologist; 3) an initial pilot/training study preceding the main study was to be performed in two groups of 7- and 9-year-old children, and results were discussed among participants during a workshop; 4) test protocols of the Bender test (GFT version) were distributed by the coordinating institute (CI), and feedback concerning scoring discrepancies was given; 5) videotapes of test sessions were sent to CI by some participants and discussed at a workshop; 6) site visits of CI delegates were arranged to discuss and solve difficulties, and to clarify procedural aspects.

Statistical Model and Analysis

The full set of data collected by the participating groups was available for statistical treatment at the CI. Multiple regression analysis was the statistical tool in testing hypotheses about the impact of environmental lead exposure on neuropsychological development. Apart from blood lead as the independent and psychological outcome as the dependent variable, the full model included age and gender as extraneous variables and father's occupational status as well as maternal education as social confounders. The final decision for these two covariates to serve as surrogates for social background was based on the following considerations: Cross-cultural differences necessitated a simplified approach. Since paternal and maternal education were highly correlated, and since, furthermore, maternal education was the more complete data set, mother's education alone was finally selected to cover this aspect of social background. Some details of definition are given in Table 2. The highly diverse occupational systems proposed by the participating groups induced us to simply define a skilled/unskilled dichotomy to cover father's occupational status. It should be pointed out here that some participants had used more complex confounder models in their already published individual analyses [e.g., (3,9)].

RESULTS

Characteristics of Individual Studies

Summary information about the individual studies in terms of sample size, blood-lead concentrations (PbB) and tooth-lead concentrations (PbT) and study type is given in Table 1. Studies differ markedly in terms of sample size as well as in terms of exposure. This latter difference is not simply related to study type, because both high- and low-average PbBs occur within the lead smelter category, whereas general population studies cover both intermediate and high-average exposures. Although blood-lead and tooth-lead values did correlate moderately (.31–.81) in all but one study, there were, nevertheless, some surprising discrepancies in terms of PbB/PbT quotients (e.g., in groups 3, 5 and 6). This, in addition to the small number of teeth, was another reason to discard tooth-lead levels as markers of environmental exposure for subsequent regression analyses.

TABLE 1

CHARACTERIZATION OF INDIVIDUAL STUDIES IN TERMS OF BLOOD-LEAD (µg/dl) AND TOOTH-LEAD LEVELS (µg/g), SAMPLE SIZE AND STUDY TYPE

Study	Blood-Lead n	Blood-Lead x_g	Blood-Lead s_g	Tooth-Lead n	Tooth-Lead x_g	Tooth-Lead s_g	Study Type
1	460	22.0	1.4	—	—	—	Lead Smelter
2	301	18.9	1.3	—	—	—	General Population
3	254	18.2	1.7	68	2.6	1.8	General Population
4	216	11.0	1.3	116	6.1	1.6	General Population
5	142	18.2	1.6	71	8.1	1.7	General Population
6	48	18.0	1.5	48	14.8	2.1	Lead Smelter
7	109	8.3	1.4	77	4.5	1.7	Lead Smelter
8	109	7.4	1.3	52	3.6	1.8	Lead Smelter

Geometric means (x_g) and geometric standard deviations (s_g) are given.

Summary information about the individual studies in terms of covariates is given in Table 2. This table, furthermore, briefly summarizes the definition of maternal education as proposed by the participating groups. Whereas there was satisfactory comparability of studies in terms of age and gender, social indicators exhibited pronounced intergroup variability.

Neurobehavioral Outcome

Instead of looking at statistical significance alone, consistency of exposure-related outcome across studies will be taken into account as well. Table 3 summarizes *t*-values indicating the statistical significance of the regression coefficients for the different outcome measures as found in the individual studies.

Significant associations between blood-lead values and the different behavioral endpoints are restricted to groups 1 and 7. Rather consistent results in terms of directionality exist for both versions of the Bender test (GFT): All but one study (6) reported

TABLE 2

COMPARISON OF STUDIES IN TERMS OF EXTRANEOUS VARIABLES AND COVARIATES

Study	Gender % Female	Age (years) Mean	Age (years) SD	Paternal Occupation % Unskilled	Maternal Education % Lowest Category
1	46.5	9.5	1.4	36.9	7.7[a]
2	48.5	9.2	1.4	1.4	5.7[b]
3	51.2	8.5	0.6	61.7	—[c]
4	45.0	7.8	0.4	19.6	40.0[d]
5	49.3	7.3	0.4	53.3	17.5[e]
6	42.9	6.7	0.7	12.2	—[c]
7	49.1	6.5	0.3	18.4	30.7[f]
8	43.2	8.3	0.6	18.0	32.7[f]

[a]Years of schooling; 3 categories.
[b]Level of education; 6 categories.
[c]Not available.
[d]Years of schooling and level of education; 4 categories.
[e]Levels of schooling; 3 categories.
[f]Occupational status and level of education; 3 categories.

TABLE 3

RESULTS OF INDIVIDUAL STUDIES IN TERMS OF t-VALUES FROM MULTIPLE REGRESSION ANALYSIS DESCRIBING THE DEGREE OF ASSOCIATION BETWEEN BLOOD-LEAD LEVELS AND DIFFERENT BEHAVIORAL ENDPOINTS

Behavioral Endpoints		Groups 1	2	3	4	5	6	7	8
WISC		−3.14‡	−0.14	0.42	0.21	0.95	−0.33	−1.03	−0.76
GFT	N	1.96†	1.77	0.03	0.50	1.14	−2.01	0.78	0.92
	I	2.29†	0.84	0.34	0.31	1.43	−0.83	—	1.47
Trail-Making		0.40	−1.00	−0.50	1.33	0.0	—	—	—
Vienna Hits	Slow	−3.93§	—	0.98	—	—	—	−1.81*	−0.89
	Fast	−3.87§	—	0.76	—	—	—	−1.96*	0.13
Vienna Errors	Slow	4.06§	—	−2.23	—	—	—	2.19†	0.51
	Fast	2.93‡	—	−1.15	—	—	—	5.16§	1.72*
DRT		1.81	—	−1.22	0.39	—	—	—	—
Needleman Teacher		—	−0.27	−0.14	0.13	−0.29	1.66	—	—
Needleman Parents		1.29	1.18	−1.29	−0.50	−1.75	−0.92	0.64	−0.96

*$p<0.10$; †$p<0.05$; ‡$p<0.01$; §$p<0.001$ (one-tailed p-values).

positive associations between lead exposure and the GFT error score, although significant for group 1 only. Disruption of serial choice reaction performance as assessed by hits and errors of the Vienna Reaction Device (VRD) occurs in three of the four studies considered here; two studies had to be discarded for this behavioral paradigm for methodological uncertainties. Inconsistent results in terms of directionality occur for the remaining variables. This may best be seen for those outcome measures implemented in all eight studies, namely the WISC and the parental behavior ratings by means of the Needleman scales. Positive and negative associations with blood lead were produced here.

In order to illustrate dose-response characteristics for some of those variables for which almost complete information was available from all eight studies, results from the WISC and from both GFT versions were plotted against blood-lead concentrations (Figs. 1, 2 and 3). For reasons of comparability across studies and for descriptive purposes, blood-lead concentrations (PbB) were grouped into classes of equal logarithmic width in all these figures, with PbB values (x-axis) representing geometric means of these classes. Figure 1 depicts WISC results (mean of four subtests) from all the individual studies. The between-groups variability is noteworthy and is larger here than for any of the other endpoints. These systematic differences are mainly due to inappropriate age norms of the old WISC as opposed to the revised form (WISC-R), as well as to a lack of culturally adapted normative data in some of the participating countries. Apart from such systematic differences the within-groups variability is very small indeed. Except for group 1, no systematic decline of WISC standard scores with increasing blood-lead concentrations can be inferred from Fig. 1.

The between-groups variability is less pronounced for GFT performance in terms of the error scores of versions N (Fig. 2) and I (Fig. 3). In addition, there is a trend for errors to increase with increasing blood lead in most of the individual studies; it is clearly impossible, however, to locate a critical PbB level which might be identified as an effect threshold. This is true for both versions of the Bender test.

Despite existing study differences, both in terms of intercept and slope of individual dose-response functions, as is apparent from Figs. 1–3, an effort was made to evaluate the overall effects of lead exposure on the different neurobehavioral endpoints using the full set of data and, after having considered social confounding for each study separately, treating individual studies as dummy variables. The results of this overall regression analyses are summarized in Table 4. This table contains the following message. 1) There is a weak negative association for environmental lead exposure and psychometric intelligence which is of only borderline significance. 2) Stronger and more consistent associations with lead exposure exist for disruption of visual-motor integration (Bender test, both versions) and of serial choice reaction performance (VRD) in terms of hits and errors. 3) No significant effects exist for less standardized measures of neurobehavioral functions (Trail-Making test) and for crude behavior ratings by both parents and teachers. 4) Even if significant, the degree of association with

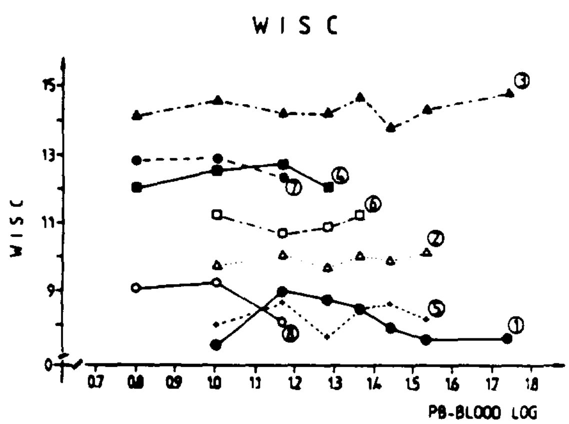

FIG. 1. Dose-response information for individual studies with WISC results (mean of four subtests) as the behavioral endpoint. The abscissa is log PbB (μg/dl).

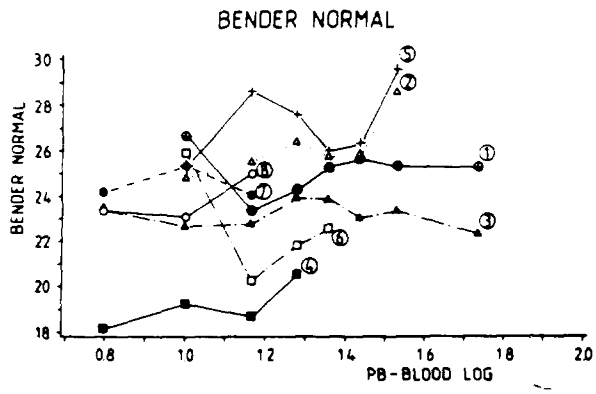

FIG. 2. Dose-response information for individual studies with GFT results (error score of normal version) as the behavioral endpoint. The abscissa is log PbB (μg/dl).

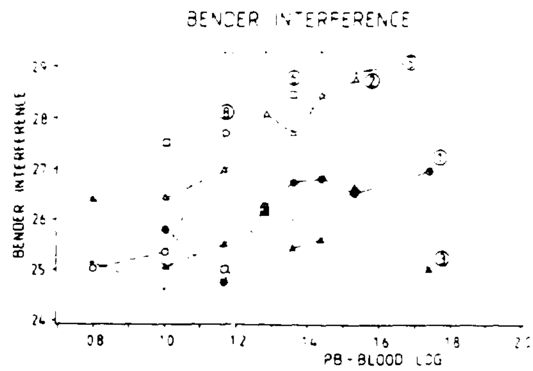

FIG. 3. Dose-response information for individual studies with GFT results (error score of interference version) as the behavioral endpoint. The abscissa is log PbB (μg/dl).

blood-lead concentration is small and never exceeds 0.8% of the total variance.

DISCUSSION

Neurobehavioral effects of environmental lead exposure in children represent weak signals embedded in a noisy background. Many of the published cross-sectional studies suffer from insufficient power to extract subtle effects (8). This may be due to inadequate methodology, insufficient sample size, narrow range of exposures, or a combination of these shortcomings. Combining the results of several such studies by means of quantitative metaanalysis has been used to arrive at a more valid estimate of the "true" effect size (8).

This approach suffers from the following inherent weaknesses. 1) Reliance on published information often introduces a bias, since papers with insignificant findings tend to remain unpublished. 2) The studies considered for metaanalysis are necessarily heterogeneous in terms of general design, markers of lead exposure, confounder structure, as well as overall statistical analysis. 3) This approach is necessarily based on selective information provided by the authors, rather than on independent analysis of the original data.

The results of the present WHO/CEC cross-sectional study may be taken to represent realistic risk assessment for environmental lead exposure in children, because original data collected in individual studies with a common study protocol and inherent quality control elements were subjected to uniform statistical analysis, and because findings were extracted from a large sample of about 1800 children and based on a broad range of exposure values from below 5 to about 60 μg Pb/100 ml blood.

Within this range the study confirmed that there are detectable

TABLE 4

SUMMARY OF RESULTS BASED ON OVERALL MULTIPLE REGRESSION ANALYSIS OF THE FULL SET OF DATA ACROSS STUDIES

Behavioral Endpoint		No. of Labs	No. of Children	Slope b	Significance t	PbB $r^2 \times 100$
WISC		8	1698	−0.53	−1.62*	.19%
GFT	N	8	1698	1.64	2.17†	.31%
	I	7	1584	2.03	2.54‡	.46%
Trail-Making		6	1433	0.00	0.11	.03%
Vienna Hits	Slow	4	971	−16.11	−2.57‡	.76%
	Fast	4	971	−11.51	−2.39‡	.67%
Vienna Errors[1]	Slow	4	971	0.15	1.68†	.37%
	Fast	4	971	0.20	2.37‡	.66%
Delayed RT[2]		3	986	0.01	0.41	.02%
Needleman Teacher[3]		5	1012	−0.00	−0.03	.00%
Needleman Parents[3]		8	1698	−0.03	−0.67	.00%

[1] x' = log (x + 0.5); [2] log (msec); [3] sum of 6 items.
*$p<0.1$; †$p<0.05$; ‡$p<0.01$ (one-tailed p).

TABLE 5

RISK ASSESSMENT FOR ENVIRONMENTAL LEAD EXPOSURE IN TERMS OF BLOOD-LEAD CONCENTRATIONS (PbB) FOR PSYCHOMETRIC INTELLIGENCE (WISC) AND VISUAL-MOTOR INTEGRATION (GFT NORMAL AND INTERFERENCE VERSIONS)

PbB	WISC			GFT-N			GFT-I		
		% Below*			% Above*			% Above*	
μg/dl	Mean	1 SD	2 SD	Mean	1 SD	2 SD	Mean	1 SD	2 SD
5.0	109.6	14.0	1.9	22.5	12.7	1.6	23.6	11.9	1.5
7.9	108.5	14.7	2.0	22.9	14.0	1.8	24.0	13.4	1.7
12.6	107.5	15.4	2.2	23.2	15.2	2.1	24.4	14.9	2.1
20.0	106.4	16.4	2.4	23.5	16.6	2.4	24.8	16.6	2.4
31.6	105.4	17.1	2.6	23.8	18.1	2.7	25.2	18.7	2.9
50.1	104.3	17.9	2.7	24.2	19.5	3.1	25.7	20.6	3.4

*Proportion of children in two risk categories (1 and 2 SD) is given.

exposure-related neurobehavioral effects in school-age children. The strongest and most consistent effects were observed in established clinical tests of visual-motor integration, namely both GFT versions, as well as in serial choice reaction performance, namely the Vienna Reaction Device. The degree of association was significant or highly significant for these variables, although the variance explained by PbB (r^2) never exceeded 0.8%. Psychometric intelligence, as assessed by means of the WISC, was also affected by lead exposure, although the effects were inconsistent across studies and the overall degree of association was of only boderline significance. In all of these cases higher lead levels were associated with poorer performance. No significant associations with blood lead were found for less standardized measures of visual-motor integration, namely the Trail-Making test, and for ratings of general behavior by parents and teachers.

In order to transform the statistical parameters given in Table 4 into estimates of risk, the values given in Table 5 were calculated for the two GFT versions and the WISC as measures of visual-motor integration and of cognitive capacity, respectively. These behavioral endpoints were selected for risk assessment, because they are based on the most comprehensive data set within this study.

Using averages and observed standard deviations (SD), the percentage of children exceeding statistical criteria of deficit, namely 1 or 2 SD above or below the respective mean values, was calculated by means of the regression coefficients given in Table 4. It may, thus, be seen, for example, that by increasing blood-lead levels from a low 5 μg/dl to an elevated level of 20 μg/dl, the proportion of children in the intermediate risk category (1 SD) increases by 2.4% (WISC), 3.9% (GFT-N) or 4.7% (GFT-I).

An important additional observation of this study was that the full data set did not allow us to identify an effect threshold, that is, a biological exposure limit above which adverse health effects in children are to be expected (see Figs. 1, 2 and 3). In the combined analysis of the full set of data across studies, a significant linear relationship was found between blood-lead concentration and visual-motor integration as well as serial choice reaction performance. The spectrum of Pb effects was, thus, largely consistent with what has been found in previous cross-sectional research (12), although the poor consistency of Pb-related IQ deficit across studies and its small effect size does not fit smoothly into the overall pattern of findings, as evaluated in a recent metaanalysis of published studies (8).

ACKNOWLEDGEMENTS

This study was partly supported by grant No. 116 06 076 of the Umweltbundesamt (Berlin). Further support was available from WHO/EURO (Copenhagen) and CEC (Brussels); we would like to thank Dr. Dinko Kello (WHO/EURO) and Dr. Andrew Sors (CEC) for continued interest and help. Thanks are due, in particular, to our research colleagues of the different participating groups for excellent cooperation and fruitful discussions. We are, furthermore, grateful to Prof. Raymond Cluydts (Brussels) and to Prof. William Yule (London) for technical support and continued scientific advice.

REFERENCES

1. Bender, L. A visual-motor Gestalt test and its clinical use. New York: American Orthopsychiatry Association, No. 3; 1938.
2. Canter, A. A background interference procedure to increase the sensitivity of the Bender Gestalt test to organic brain disorder. J. Consult. Psychol. 30:91–97; 1966.
3. Hatzakis, A.; Kokkevi, A.; Maravelias, C.; Katsouyani, K.; Salaminos, F.; Kalandidi, A.; Koutselinis, A.; Stefanis, C.; Trichopoulos, D. In: Smith, M.; Gratn, L. D.; Sors, A. I., eds. Lead exposure and child development. An international assessment. Dordrecht: Kluwer Academic Publ.; 1989:211–223.
4. Hunter, J.; Urbanowicz, M. A.; Yule, W.; Lansdown, R. Automated testing of reaction time and its association with lead in children. Int. Arch. Occup. Environ. Health 57:27–34; 1985.
5. Klebelsberg, D. Wiener Determinationsgerät. Diagnostika VI; 1960.
6. Lansdown, R. Lead, intelligence, attainment and behaviour. In: Lansdown, R.; Yule, W., eds. The lead debate: The environment, toxicology and child health. London: Croom Helm; 1986:235–270.
7. Needleman, H. L.; Gunnoe, C.; Leviton, A.; Reed, R.; Peresie, H.; Maher, C.; Barret, P. Deficits in psychological and classroom performance in children with elevated dentine lead levels. N. Engl. J. Med. 300:689–695; 1979.
8. Needleman, H. L.; Gatsonis, C. A. Low-level lead exposure and the IQ of children. A meta-analysis of modern studies. JAMA 263: 673–678; 1990.
9. Nörby-Hansen, O.; Trillingsgaard, A.; Beese, I.; Lyngbye, T.; Grandjean, P. Neuropsychological profile of children in relation to dentine lead level and socioeconomic group. In: Smith, M.; Grant, L. D.; Sors, A. I., eds. Lead exposure and child development. An international assessment. Dordrecht: Kluwer Academic Publ., 1989: 240–250.
10. Reitan, R. Validity of the trail making test as an indicator of organic brain damage. Percept. Mot. Skills 8:271–276; 1958.
11. Schlange, H.; Stein, B.; Bötticher, J.; Taneli, S. Göttinger Formreproduktionstest (GFT). Göttingen: Hogrefe; 1972.
12. Smith, M.; Grant, L. D.; Sors, A. I., eds. Lead exposure and child development. An international assessment. Dordrecht: Kluwer Aca-

demic Publ., 1989.
13. Umweltbundesamt, ed. Studie zur Verhaltenstoxizitat von Blei bei Kindern. WHO/CEC-Projekt, in preparation, 1990.
14. U.S. Environmental Protection Agency (EPA), ed. Air quality criteria for lead, vol. IV. Research Triangle Park: USEPA Environmental Criteria and Assessment Office; 1986.
15. Vivoli, G., Bergomi, M., Borella, P.; Fantuzzi, G.; Simoni, L.; Catelli, D.; Sturloni, N.; Cavazzuti, G. B.; Montorsi, R.; Campagna, R.; Tampieri, A.; Tartoni, P. L. Evaluation of different biological indicators of lead exposure related to neuropsychological effects in children. In: Smith, M.; Grant, L. D.; Sors, A. I., eds. Lead exposure and child development. An international assessment. Dordrecht: Kluwer Academic Publ.; 1989:224–239.
16. World Health Organization, ed. Joint WHO/CEC Final Review Workshop on the Lead Neurotoxicity Study in Children. Technical Report, in preparation, 1990.
17. Winneke, G.; Collet, W.; Kramer, U.; Brockhaus, A.; Ewert, T.; Krause, C. Follow-up studies in lead-exposed children. In: Smith, M.; Grant, L. D.; Sors, A. I., eds. Lead exposure and child development. An international assessment. Dordrecht: Kluwer Academic Publ.; 1989:260–270.
18. Winneke, G.; Brockhaus, A.; Collet, W.; Kramer, U. Modulation of lead induced performance deficit in children by varying signal rate in a serial choice reaction task. Neurotoxicol. Teratol. 11:587–592; 1989.
19. Wittling, W. W. Neuropsychologische Diagnostik. In: Groffmann, K. J.; Michel, L., eds. Verhaltensdiagnostik. Göttingen: Hogrefe; 1983:193–335.